Structural Equation Modeling

Mgmt 290
Lecture 1
Sep 26, 2005

Why SEM?
Use of SEM in IS Journals

- 18% in mid-1990s and growing

	Information &Management	Information System Research	MIS Quarterly	All Three
Total	8%	45%	25%	18%

- In "SEM and Regression – Guidelines for Research Practice" By D Gefen, D.W. Staub, and M. Boudreau, Communication of AIS Vol. 4 Article 7- August 2000
- A lot of SEM application in management, sociology, psychology, and even in political science now.

The Goals of the Course

- To have students able to construct, analyze, modify, estimate, evaluate and explain structural equation models and report the results in a manner acceptable in professional journals.

About the Instructor – Alex Liu

M.S. of Statistics and Ph.D. of Sociology from Stanford

Research on Methodology and Consumer Behavior

Practical Experience in Entrepreneurship and Business Innovation

Worked for Hoover Institution & Asia/Pacific Research Center at Stanford and A Few Business Groups

Now mainly conduct Research Methods Studies under the RM Institute, plus consulting for organizations like UN and IBM Research

My Methods Approach

- RM = Statistical Techniques + Statistical Strategies + Subject Knowledge.
- Research is a step by step process.
- My Main Project: Using RM4Es to represent a research status and ResearchMap to represent a research process, that leads to intelligent assistance for empirical research.
- More in www.ResearchMethods.org

My Work on SEM

- SEM for Causal Analysis (Use SEM evidence to infer causal relationship)
- Exploratory SEM.
- SEM to Model Consumer Behavior.
- Latent Variable Model for evaluation.

Main Content of This Course

- What will be covered:

- Path Analysis (structural model)
- Confirmative Factor Analysis (measurement model)
- Hybrid Model (structural and measurement)
- SEM Software - LISREL (SPSS)

- SEM Model Building Process (Stat Strategies + how to combine SEM with your subject knowledge for research)

- (Not multiple level SEM)

Focus & Style of the Course

- Focus on application to research
- Combine LISREL and SEM model for YOUR Research Project

- Active Participation is Needed (first 2/3 with more on lectures and demos, and the last 1/3 with more discussion)
- Learn the basics, and go through a research process together

Readings for the course

- **Rex B. Kline 2004 Principles and Practise of Structural Equation Modeling, Guilford Press**
- **E. Kevin Kelloway 1998 Using LISREL for Structural Equation Modeling, Sage Publications**
- Otis Dudley Duncan 1975 Introduction to Structural Equation Models, New York: Academic Press
- Ken Bollen 1989 Structural Equation Modeling with Latent Variables, Wiley
- Gerhard Mels 2003 Getting Started with the Student Edition of LISREL 8.53 for Windows, Scientific Software International

Requirements

- **Understand the basics of the SEM structure model and measurement model**
- **Learn to use LISREL**
- **Use a Dataset (your own or D&D data)**
- **Complete 4 Assignments**
- **Complete a Final Research Presentation (research paper is optional)**

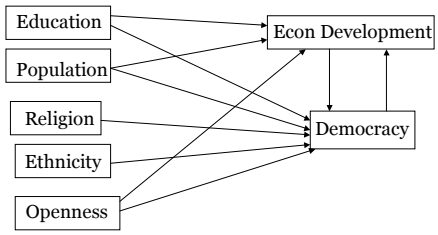
Assignments

- Assignment One - Research Proposal (due in the third week in class)
- Assignment Two - Initial Model Specification (due in the fifth week in class)
- Assignment Three - Model Estimation (due in the seventh week in class)
- Assignment Four - Model Modification and Final Model (due in the ninth week in class)
- Final Presentation in the last week

About the dataset

- Your own dataset
- Or the D & D (development & democracy) data
- economic development and political development from 1950 ~ 1990 (2000)
- http://www.researchmethods.org/sem-data.htm

Possible Research from Using the D&D Data



Course Web Pages

- Please Visit
- www.ResearchMethods.org/sem.htm
- For the details of administrative issues
- (outline, lecture notes, data, …)
- AND for updates

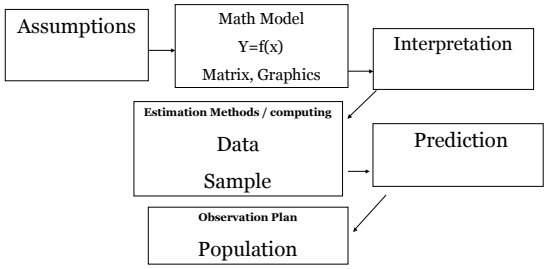
Questions about the course?

General Pre-requisites

- Linear Regression Analysis
- Probability, estimation and hypothesis testing
- Basics of factor analysis
- Basics of measurements (reliability and validity)

Quantitative Research Logic

- base for applying SEM in your research



The research process - to use SEM properly

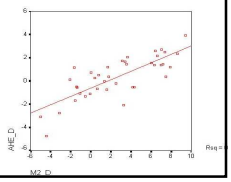
- Test a hypothesis (significance testing)
- - confirmative process
- Specify -> estimate -> modify – estimate ....
- - exploratory process

Some Statistics Concepts

- The good understanding of the following concepts is required:
  - Measurement levels (nominal, ordinal, interval and ratio)
  - Correlation and covariance
  - Indicators (reliability and validity)
  - Goodness fit indicator (R^2 ...)
- 4Es (Equation, Estimation, Errors, Explanation) to summarize a research

More about correlation and covariance

- Correlation $r_{XY} = \Sigma Z_x Z_y / (N-1)$
- Covariance $Cov_{XY} = r_{XY} SD_X SD_Y$
- (use scatter plot to visualize correlation)
- $B = r_{XY} SD_Y / SD_X$
- $Z_x = (X - mean) / SD_X$



Partial Correlation

- $r_{XY.W} = (r_{XY} - r_{XW} r_{YW}) / Sqrt[(1-r_{XW}^2)(1-r_{YW}^2)]$

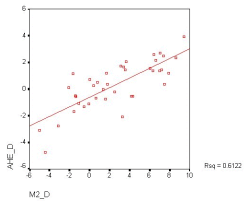
Partial regression plot (added variable plot)

1) Fit regression of Y on all Xs except X_k, save the residuals as $e_{Y(Xk)}$
2) Fit the regression of X_k on all the other Xs, save the residuals as e_k
3) Plot $e_{Y(Xk)}$ vs e_k

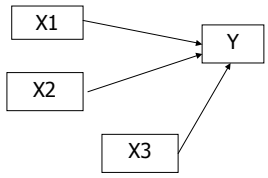
Simple Regression Model

- $Y_i = \beta_o + \beta_1 X_i + \varepsilon_i$

- ***Constant***
- ***(intercept)***
- ***Slope***
- ***residuals***



Extended to Multiple Regression



- $Y_i = \beta_o + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$
- A special case of SEM (one equation vs. many equations)

Assumptions for regression

- Variables measured without error and at the interval level
- $E(e_j)=0$ --- the mean value of the error term is 0
- $Var(e_j) = \sigma^2$ --- the variance of the error term is constant
- $Cov(e_i, e_j) = 0$, no autocorrelation
- No serious collinearity
- E_j is normally distributed
- Linear and additive between ind vars and dep var

More Issues to Deal in Regression Modeling

- Missing Values
- Outliners
- Model Misspecification (omitted variables)

SPSS Results

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.034[a]	.001	.000	.77

a. Predictors: (Constant), rate pol sys today

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.013	1	1.013	1.695	.193[a]
	Residual	894.781	1498	.597		
	Total	895.794	1499			

a. Predictors: (Constant), rate pol sys today

b. Dependent Variable: confidence:legal system

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.715	.022		122.268	.000
	rate pol sys today	-1.83E-03	.001	-.034	-1.302	.193

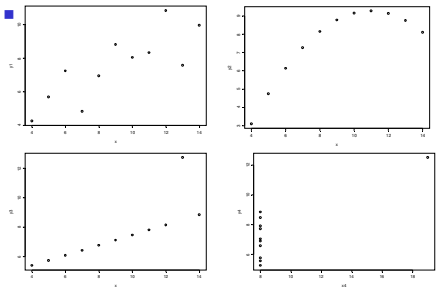
a. Dependent Variable: confidence:legal system

The Danger of Simply Applying OLS

	X1-3	Y1	Y2	Y3	X4	Y4
1	10.0	8.04	9.14	7.46	8.0	6.58
2	8.0	6.95	8.14	6.77	8.0	5.76
3	13.0	7.58	8.74	12.74	8.0	7.71
4	9.0	8.81	8.77	7.11	8.0	8.84
5	11.0	8.33	9.26	7.81	8.0	8.47
6	14.0	9.96	8.10	8.84	8.0	7.04
7	6.0	7.24	6.13	6.08	8.0	5.25
8	4.0	4.26	3.10	5.39	19.0	12.50
9	12.0	10.84	9.13	8.15	8.0	5.56
10	7.0	4.82	7.26	6.42	8.0	7.91
11	5.0	5.68	4.74	5.73	8.0	6.89

B_0=3.00
B_1=0.50
RSS=13.75

The Plots for the four datasets



Modeling Process Example: A 7 Step Approach

- Step 1: Data Cleaning & First OLS Results – missing data, added variable plots, dummy variables, collinearity, first OLS estimation
- Step 2: Diagnostics 1 – normality, homoscedasticity, autocorrelation (Q-Q plot, Residual Plot, Durbin-watson)
- Step 3: Diagnostics 2 – linearity and outliners (residual plot and studentized residuals)
- Step 4: Variable Transformation

My 7 Step Approach to Build a Regression Model - 2

- Step 5: Model Assessment & Validation – variable selection & model validation (step-wise and cross-validation)
- Step 6: Diagnostics Again – if problems go back to step 4
- Step 7: Final OLS Estimates – if necessary, use non-OLS methods

Example: Political Confidence

- Dependent Variable: PolConf – General measure of political confidence on legal system, federal government, pol parties, parliament – 4 ~ 12

- 7 Independent Variables:
- Left ~ Right (1 left … 10 Right)
- Church (church attendance 1 more than once a week … 7 practically never)
- Age
- Education
- Income
- Postmaterialism/materialism (0 ~ 5 post)
- Marriage status (1 married, 2 living together, 3 divorced, 4 separated, 5 widowed, 6 single)

Example for Step 1

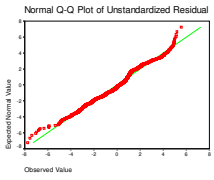
Coefficients[a]

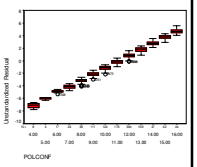
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10.325	.478		21.618	.000
	Mat/Postmat (12-item index)	.114	.059	.061	1.944	.052
	income	5.498E-02	.033	.059	1.690	.091
	marital status	4.249E-03	.038	.004	.113	.910
	left-right self placement	-3.05E-02	.036	-.026	-.843	.400
	Church Attendance	.118	.034	.108	3.501	.000
	age	4.228E-03	.004	.032	1.041	.298
	education- codes vary cross-nationally	-3.65E-02	.035	-.035	-1.053	.292

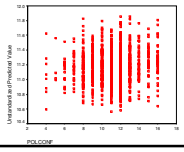
a. Dependent Variable: POLCONF

Example for Step 2

- Durbin-watson = 1.822







Example for Step 3

- Identified about 6% Outliers by Using Studentized Residuals & Deleted them (study them separately is also recommended)
- No collinearity Problem as
- VIF < 4
- Condition Index < 18

- Recode Married into
- a Dummy Variable
- 1 – married or
- living together
- 0 - others

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10.790	.522		20.688	.000
	marital status	3.54E-02	.485	-.009	-.073	.942
	left-right self placeme	-4.48E-03	.031	-.005	-.143	.886
	age	1.740E-03	.003	.044	1.354	.176
	education- codes vary cross-nationally	-.156	.051	-.180	-3.049	.002
	income	1.405E-03	.028	.011	.297	.767
	Mat/Postmat (12-item index)	.261	.086	.167	3.028	.003
	Church Attendance	.151	.053	.164	2.867	.004
	MPOST	-.127	.105	-.098	-1.210	.227
	MCHURCH	-.145	.063	-.180	-2.319	.021
	MEDU	.126	.060	.236	2.097	.036

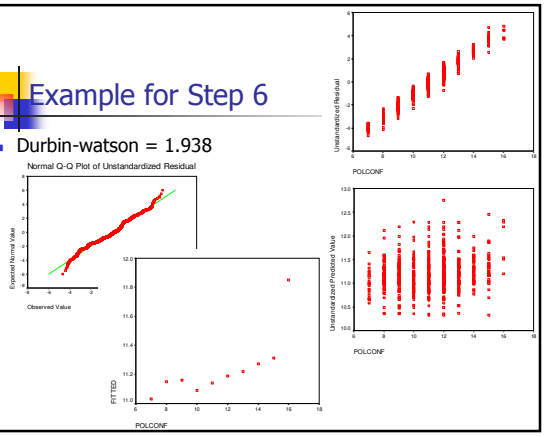
a. Dependent Variable: POLCONF

Example for Step 5

- Use Backward Method
- Church, PostMaterialism, Edu, Medu, Mchurch,
- R^2 = .168

Example for Step 6

- Durbin-watson = 1.938



Step 7: Final OLS Estimates

Present & Interpret
Your Results Now

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	11.220	.384		29.231	.000
	marital status	-.268	.461	-.066	-.582	.561
	education- codes vary cross-nationally	-.149	.049	-.172	-3.070	.002
	Mat/Postmat (12-item index)	.172	.049	.110	3.477	.001
	Church Attendance	.148	.051	.161	2.875	.004
	MCHURCH	-.144	.062	-.179	-2.335	.020
	MEDU	.116	.058	.218	2.002	.046

a. Dependent Variable: POLCONF

Final Product of This Course Research Paper or Presentation

- Introduction
- Argument
- Measurements/Variables/Dataset
- Methods
- Results
- Discussion
- Conclusion